Results and Conjectures in Mathematical Relativity

Robert Bartnik

In 1916 [1] Albert Einstein presented his well-known equation

$$R_{\alpha\beta} - \frac{1}{2} R\, g_{\alpha\beta} = 8\pi\, T_{\alpha\beta}\,,$$

relating the geometry of space and time, modelled by a 4-dimensional manifold with Lorentz metric $g_{\alpha\beta}$ and Ricci curvature $R_{\alpha\beta}$, to the physical matter distribution, modelled by the stress-energy tensor $T_{\alpha\beta}$. This was a truly revolutionary theory in that it lead to major changes, both philosophical and scientific, in the way we view our world.

The effect on physics has perhaps been the most noticeable. Quite apart from its surprising predictions about the large scale geometry of the universe, general relativity introduced the idea of a "Theory Of Everything", now a Holy Grail of theoretical physics, and it showed that differential geometry is the natural language of physics. The success of the Yang–Mills–Higgs model of the electroweak interaction, and of the intensive work in areas such as string theory and supergravity, shows clearly that these ideas have had some impact on physics.

The effect of Einstein's theory on mathematics, especially differential geometry, was just as pronounced. By emphasising the physical relevance of such questions as the relation between curvature, geodesics and geometry, and the nature of Maxwell's equations on a manifold, the outlines of the subject we now recognise as differential geometry were laid.

However, it can be said that it is not the Einstein equations themselves, but rather the paradigm they represent – that the techniques and results of geometry are of fundamental importance on physics – that has been more significant. The Einstein equations have been comparatively neglected, and the reasons for this are not hard to find. The physical effects predicted by the theory are extremely small and thus difficult to detect, and it is not possible to arrange experiments – the universe is our only laboratory. Mathematically the (vacuum, $T_{\alpha\beta} = 0$) Einstein equations form a non-linear system of hyperbolic partial differential equations with a coordinate gauge degeneracy, and even now our understanding of such systems is very limited. The only hope of getting information about physically interesting situations, such as the 2–body problem, is to resort to numerical computation, or, in nearly Newtonian situations, to use approximation methods.

Despite all this, there is beautiful and even surprising mathematics hidden in general relativity, and I would like to describe some examples of this here and to indicate some open problems. A recurring theme will be the way physical intuition and mathematical theorems complement each other: physics suggests mathematical conjectures, and mathematical theorems in turn validate and enhance physical intuition. The most spectacular example of this is the positive mass theorem [2,3], but there are other examples. Presumably this correspondence between the mathematics and the physics just illustrates that general relativity is a physically realistic theory, but I still find it rather mysterious.

The Schwarzschild metric

A large proportion of the physical intuition in general relativity can be traced back to one special solution, the Schwarzschild spacetime $\mathcal{S}$, which was also uncovered in 1916 [4]. We quickly summarise its basic properties. The metric of $\mathcal{S}$ in the usual coordinate chart ($r > 2M$, $t \in \mathbb{R}$, $\Omega \in S^2$) is

$$ds^2_{\text{Schw.}} = -\left(1-\frac{2M}{r}\right) dt^2 + \left(1-\frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{1}$$

where M is a positive constant and $d\Omega^2$ is the usual metric on S^2. In fact, this expression contains two solutions of the vacuum Einstein equations, represented by the coordinate ranges $0 < r < 2M$ and $2M < r$, and it was realised quite early that these covered two disjoint regions in a larger vacuum spacetime (also denoted $\mathcal{S}$), with the surface $\{r = 2M,\ t = +\infty\}$ representing their common boundary. This surface is called the *horizon* of $\mathcal{S}$ and separates the (black hole) region $\{r < 2M\}$, where every observer (observer = future–directed, time–like curve) must hit the curvature singularity at $r = 0$ in finite proper time, from the exterior region $\{r > 2M\}$, where observers can avoid this presumably ghastly fate. These properties can be easily shown by introducing the Kruskal–Szekeres coordinates, see [5].

Some further properties of $\mathcal{S}$ which will be important later are

(a) *time–independence*: the metric ds^2 has a time like Killing vector $K = \partial_t$ which is also hypersurface–orthogonal. Such metrics are called *static*, and more generally, a metric with a timelike Killing vector is called *stationary*.

(b) *spherical symmetry*: the group $SO(3)$ acts isometrically on $\mathcal{S}$ with orbits diffeomorphic to S^2.

(c) *asymptotically flatness* (AF): introducing rectangular coordinates (x^i) into the region $\{r > 2M\}$ in any obvious way, we find that the spatial metric $g_{ij}(x)$ in these coordinates satisfies

$$g_{ij}(x) - \delta_{ij} = O(|x|^{-1}) \quad ,$$

$$|\partial_k g_{ij}| = O(|x|^{-2}) \quad , \quad |\partial_k \partial_l g_{ij}| = O(|x|^{-3}) \quad , \quad \text{as } |x| \to \infty .$$

(d) *mass*: a timelike geodesic $s \longrightarrow (r(s),t(s))$ in $\mathcal{S}$ satisfies

$$\frac{d^2r}{ds^2} = -\frac{M}{r^2} + O(r^{-3}) \quad , \quad r >> 2M \quad ,$$

where s is the proper time parameter. Using the AF coordinates of (c) to make a comparison with Newtonian gravity, we are lead to identify M as the mass of $\mathcal{S}$.

We must also mention two generalisations of $\mathcal{S}$ which will be needed: the Reissner–Nordström spacetime [6] and the Schwarzschild static star [4] . The Reissner–Nordström metric is

$$ds^2 = -\left(1 - \frac{2M}{r} + \frac{e^2}{r^2}\right) dt^2 + \left(1 - \frac{2M}{r} + \frac{e^2}{r^2}\right)^{-1} dr^2 + r^2 d\Omega^2 \ ; \qquad (2)$$

a spherically symmetric solution of the Einstein–Maxwell equations with electromagnetic curvature 2–form $F = \frac{e}{r^2} dr \wedge dt$. If the constants e , M satisfy $|e| < M$ then the global structure is similar to $\mathcal{S}$ [5].

The Schwarzschild static star is produced by attaching a ball of static, constant density perfect fluid (i.e. $T_{\alpha\beta} = \mathrm{diag}(\mu,p,p,p)$ where the density μ is constant) to the region $\{r > r_0\}$, for $r_0 > \frac{9}{8} 2M$, of $\mathcal{S}$. The metric in the usual Schwarzschild coordinates is

$$ds^2 = \begin{cases} -\left\{\frac{3}{2}\sqrt{1 - \frac{2M}{r_0}} - \frac{1}{2}\sqrt{1 - \frac{2M}{r_0}\left(\frac{r}{r_0}\right)^2}\right\}^2 dt^2 + \left(1 - \frac{2M}{r_0}\left(\frac{r}{r_0}\right)^2\right)^{-1} dr^2 + \\ \qquad + r^2 d\Omega^2 , \quad \text{for } 0 < r \leq r_0 , \\ \\ -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \ , \quad \text{for } r_0 \leq r < \infty , \end{cases} \qquad (3)$$

which is C^1 across the surface of the star at $r = r_0$. This metric is clearly static, with spatial surfaces $\{t = \text{constant}\}$ diffeomorphic to $\mathbb{R}^3$ and metrically constructed by adding a spherical cap to the spatial surface of $\mathcal{S}$.

Uniqueness / Rigidity theorems

A frequent phenomenon is that of "rigidity": apparently mild conditions turn out to be satisfied by only a very few spacetimes. The prototypical result here is

Birkhoff's theorem [7]: *Suppose* V *is a spherically summetric vacuum spacetime such that the area function* r *is nondegenerate (i.e.* $dr \neq 0$*). Then* V *is isometric to a subset of* $\mathcal{S}$.

The proof is a direct computation, starting from the fact that the hypotheses allow us to construct coordinates in which the metric takes the form

$$ds^2 = -\alpha^2 dt^2 + R^2 dr^2 + r^2 d\Omega^2 , \tag{4}$$

where $\alpha = \alpha(r,t)$, $R = R(r,t)$. Writing out the vacuum Einstein equations, one finds they can be integrated using a single constant, the Schwarzschild mass. We note that a similar result is true for the spherically symmetric Einstein– Maxwell equations, leading to a uniqueness theorem for the Reissner–Nordström metrics (2). Birkhoff's theorem is interpreted physically as showing that the spherically symmetric mode of the Einstein equations carries no dynamical degrees of freedom, being governed by the single mass parameter M.

The Schwarzschild spacetime satisfies another, much more subtle, uniqueness theorem, in the class of static metrics. Since this involves a number of interesting ideas, I shall describe it in a little detail. By adapting coordinates to the spacelike hypersurfaces orthogonal to the Killing vector, the general static metric can be written as

$$ds^2 = -\alpha^2 dt^2 + g_{ij} dx^i dx^j ,$$

where $(x^i, i=1,\dots,3)$ are coordinates on a spatial 3–manifold, and $\alpha = \alpha(x)$, $g_{ij} = g_{ij}(x)$. The vacuum Einstein equations for this metric reduce to the system

$$\begin{cases} \mathrm{Ric}(g) = \alpha^{-1}\nabla^2\alpha \\ \Delta_g \alpha = 0 \end{cases} \tag{5}$$

where $\nabla^2\alpha$ is the Hessian of α in the metric g. Supplementing the AF boundary condition for g_{ij} with

$$\alpha(x) \longrightarrow 1 \quad \text{as} \quad |x| \longrightarrow \infty \quad , \tag{6}$$

we have the classical result:

Theorem. *Suppose* $(\mathcal{M},g)$ *is a complete AF Riemannian 3–manifold, and* $\alpha \in C^\infty(\mathcal{M})$ *is positive and satisfies* (6) . *If* $(\mathcal{M},g,\alpha)$ *satisfy the static vacuum equations, then* $(\mathcal{M},g) = (\mathbb{R}^3,\delta)$ *and* $\alpha \equiv 1$.

Proof: The maximum principle says α has neither maximum nor minimum in $\mathcal{M}$. Hence $\alpha \equiv 1$ and then $\mathrm{Ric}(g) \equiv 0$. Since $\mathcal{M}$ is 3–dimensional, g must be flat, and then $\mathcal{M}$ AF gives $\mathcal{M} \approx \mathbb{R}^3$. □

For the Schwarzschild solution, $\alpha = \sqrt{1-\frac{2M}{r}}$ is not constant, and the region where the system (5) is elliptic has natural boundary at the horizon $\{r = 2M\}$. This motivates the boundary condition for the main Schwarzschild uniqueness theorem.

Theorem [8]. *Suppose* $(\mathcal{M},g)$ *is an* AF *Riemannian* 3*−manifold with smooth boundary, and* $\alpha \in C^{\infty}(\mathcal{M})$ *is a positive function satisfying the boundary conditions* (6) *and*

$$\alpha = 0 \quad \text{on } \partial\mathcal{M} .$$

If $(\mathcal{M},g,\alpha)$ *satisfy the static vacuum equations* (5) *then either*

$$(\mathcal{M},g,\alpha) = (\mathbb{R}^3,\delta,1)$$
$$\text{or} \quad (\mathcal{M},g,\alpha) = \left(S^2 \times \mathbb{R} , ds^2_{Schw} , \sqrt{1-\frac{2M}{r}}\right) .$$

The elegant proof I shall sketch is due to Bunting and Masood–ul–Alam. Previous results used a rather different technique due to Israel which required that $\partial\mathcal{M}$ be connected, i.e. that $\mathcal{M}$ have only one black hole. The above theorem has the physical interpretation that it is not possible for multiple black holes to exist in static equilibrium. Note that there are static vacuum metrics which can be interpreted as representing two bodies, but these have singularities [9], as the theorem implies.

sketch proof: Since g,α are C^2 near $\partial\mathcal{M}$ and $\alpha = 0$ on $\partial\mathcal{M}$, the static equations (5) show that $\partial\mathcal{M}$ is totally geodesic, and we can construct the double manifold $(\tilde{M},\tilde{g})$ by gluing two copies of $(\mathcal{M},g)$ along $\partial\mathcal{M}$. On the respective halves $\mathcal{M}_+, \mathcal{M}_-$ we introduce the conformal metrics, $\gamma_{\pm} = (\frac{1}{2}(1\pm\alpha))^4 g$ – the resulting metric, γ say, is $C^{1,1}$ across the joining submanifold. To determine the behaviour of γ near the respective infinities of $\mathcal{M}_+, \mathcal{M}_-$, we need asymptotic expansions for g, α. These are obtained by writing the static equations in $\mathcal{M}$ in terms of the metric $\gamma_0 = \alpha^2 g$ and using AF coordinates which are harmonic in γ_0. An argument gives

$$g_{ij} = \left(1 + \frac{2m}{|x|}\right) \delta_{ij} + O(|x|^{-2})$$

$$\alpha = 1 - \frac{m}{|x|} + O(|x|^{-2}) \quad \text{as} \quad |x| \to \infty \quad ,$$

for some constant $m > 0$. Thus the infinity in $\mathcal{M}_-$ is compactified by γ, and further terms in the asymptotic expansions show that γ is at least $C^{1,1}$ across the added point. In the infinity of $\mathcal{M}_+$, γ has the expansion

$$\gamma_{ij} = \delta_{ij} + O(|x|^{-2}) ,$$

and, as will be described later, the absence of a $|x|^{-1}$ term in γ implies that γ has vanishing total mass. In summary, $(\mathcal{M} \cup \{0\}, \gamma)$ is an AF, complete $C^{1,1}$ Riemannian 3–manifold, with vanishing total mass and it is easy to verify that γ has scalar curvature identically zero. The positive mass theorem, also to be described later, implies γ is flat, and it is straightforward then to show that $(\mathcal{M}, g, \alpha)$ represents the Schwarzschild spacetime. □

There are two related results which are worthy of mention. The Schwarzschild star (3) is unique amongst static perfect fluid stellar spacetimes with constant density and having only one star [10], and it is conjectured that general static perfect fluid stars are spherically symmetric. The other result is the uniqueness of the Kerr–Newman rotating charged black hole spacetime, amongst metrics which are stationary and axially symmetric and satisfy suitable boundary conditions [11]. This is proved by a uniqueness theorem for a harmonic mapping system with boundary conditions.

A quite different type of uniqueness theorem was conjectured by S.T.Yau [12] and recently proven by J.Eschenburg [13] . This concerns spacetimes satisfying the timelike convergence condition

$$\mathrm{Ric}(n,n) \geq 0 \quad \text{for all timelike vectors } n \quad ,$$

and having a timelike *line* – a past– and future–unbounded timelike geodesic which realises the distance between any two of its points. In particular, a line has no conjugate points.

Theorem. *Suppose* V *is a globally hyperbolic spacetime which satisfies the timelike convergence condition and has a line. Then there is a Riemannian* 3*–manifold* $(\mathcal{M}, ds_M^2)$ *such that* $V = (\mathcal{M} \times \mathbb{R}, ds_M^2 - dt^2)$.

(See [5] for the definition of globally hyperbolic). Eschenburg's original proof required also that V be timelike geodesically complete – the result as stated is an improvement due to Galloway, and another variation has been given by Newman [14] . Like the motivating Cheeger–Gromoll splitting theorem of Riemannian geometry [15] , the proof is based on properties of the level sets of the Buseman functions, and follows earlier work of [16] .

The physical content of the splitting theorem can be seen by attempting to construct a counterexample. For example, in the Schwarzschild static star spacetime, we consider the sequence of timelike geodesics γ_k joining the points $(t = \pm k, x_0)$, $k \to \infty$. If the geodesics γ_k were to have a limit timelike curve, then this would be a line, contradicting the theorem. Thus there is no limiting curve; the points $x_k = \gamma_k \cap \{t = 0\}$ move out to infinity and then γ_k , k large, represents a freely falling observer who starts near the star (at $t = -k$), moves out to large radius (at $t = 0$) and then falls back to the star. In other words, gravity is an attractive force, for spacetimes satisfying the global hyperbolicity and timelike convergence conditions. This indicates, for example, that the singularity at $r = 0$ of the Schwarzschild metric with negative mass, cannot be regularised in physically reasonable way. This hint that mass is positive/gravity is attractive, for global reasons, bring us to the next topic.

Mass and the Positive Mass Theorem

The identification of the parameter M in the Schwarzschild metric with the mass of $\mathcal{S}$ was based firstly on the translation "unit–speed timelike geodesic = world–line of an observer, moving freely under the force of gravity", and secondly, on a comparison with the Newtonian space–time $\mathbb{R}^3 \times \mathbb{R}$ by means of asymptotically flat coordinates. This suggests that a definition of the mass of a more general spacetime should involve a limiting process and an AF coordinate condition. Motivated also by the Hamiltonian description of general relativity, this generalisation was found in 1962 by Arnowitt, Deser and Misner (ADM) [17]:

$$m_{ADM} = \frac{1}{32\pi} \int_{S(\infty)} (\partial_j g_{ij} - \partial_i g_{jj})\, dS^i \quad . \tag{7}$$

Here the derivatives ∂_i and metric components g_{ij} are taken in the rectangular AF (spatial) coordinates; $S(\infty)$ is the sphere at infinity, representing the limit of the integral over spheres of large coordinate radius, and dS^i is the outer normal surface measure. It is rather surprising that this (non–tensorial!) expression is independent of both the choice of limiting sequence of spheres, and of the choice of AF coordinates. Even more surprising is that the ADM mass has physical/geometric meaning:

Positive Mass Theorem. [2,3] *Suppose* $(\mathcal{M},g)$ *is a complete* AF *Riemanniann* 3*−manifold with scalar curvature* $R(g)$ *non-negative and integrable. Then* $m_{ADM} \geq 0$, *and* $m_{ADM} = 0$ *iff* $(\mathcal{M},g) = (\mathbb{R}^3,\delta)$.

The conditions of the theorem arise physically if $(\mathcal{M},g)$ is a spacelike hypersurface of a spacetime satisfying the *weak energy condition,*

$$T(n,n) \geq 0 \text{ for all timelike vectors } n$$

($T(n,n)$ is the local energy density of the observer n), and if the extrinsic curvature K_{ij} of $\mathcal{M}$ has trace zero. From the Einstein and Gauss equations we have

$$R(g) = 16\pi\, T(n,n) + \|K\|^2 - (tr_g K)^2$$

where n is the timelike unit normal vector to $\mathcal{M}$, and the condition $R(g) \geq 0$ follows. We note that the maximal surface condition, $tr_g K = 0$, implies a quasilinear elliptic equation for $\mathcal{M}$, which has been shown to have smooth spacelike solutions [18] .

There is a generalisation, the positive energy theorem [2,3] , which takes into account the contributions of the extrinsic curvature K_{ij} to the ADM energy–momentum vector, but the simpler version we have stated conveys the geometric ideas. Rather than describe the Schoen–Yau or Witten proofs, we shall give simpler proof of a weaker version [19] , which gives some insight into the geometric origin of the ADM mass.

We suppose g is an AF metric on $\mathbb{R}^3$ (or on $\mathbb{R}^n$, $n \geq 3$), which is sufficiently close to the flat metric that there are global harmonic coordinates (x^i), $i = 1,\ldots 3$ (i.e. $\Delta_g x^i = 0$), such that the metric components $g_{ij}(x)$ in these coordinates satisfy

$$|(g_{ij}(x) - \delta_{ij})\xi^i\xi^j| \leq 10^{-1} g_{ij}(x)\xi^i\xi^j \ , \text{ for all } x \in \mathbb{R}^3 , \ \xi \in \mathbb{R}^3 .$$

Now, in general coordinates the scalar curvature is

$$R(g) = \partial_i(g^{ij}(\Gamma_j - \partial_j\varphi)) - \|\nabla\varphi\|^2 + g^{ii'}g^{jj'}g^{kk'}\Gamma_{ijk}\Gamma_{i'j'k'} \ , \tag{8}$$

where Γ_{ijk} are the Christoffel symbols,

$$\Gamma_k = g^{ij}\Gamma_{ijk} = \tfrac{1}{2}g^{ij}(\partial_i g_{jk} + \partial_i g_{ik} - \partial_k g_{ij}) \ ,$$
$$\varphi = \tfrac{1}{2}\log(\det(g_{ij}))$$

and $\|\nabla\varphi\|^2 = g^{ij}\partial_i\varphi\partial_j\varphi$. The boundary term in $R(g)$,

$$\Gamma_j - \partial_j\varphi = g^{kl}(\partial_k g_{jl} - \partial_j g_{kl}) \ ,$$

is seen to converge to the ADM mass integrand, and the final term expands to

$$g^{ii'}g^{jj'}g^{kk'}\Gamma_{ijk}\Gamma_{i'k'j'} = -\tfrac{1}{4}\|\partial g\|^2 + \tfrac{1}{2}g_{kl}\partial_i g^{jk}\partial_j g^{il} \ ,$$

where $\|\partial g\|^2 = g^{ii'}g^{jj'}g^{kk'}\partial_i g_{jk}\partial_{i'}g_{j'k'}$. Integrating (8) over $\mathbb{R}^3$ with the usual Lebesgue measure dx and integrating by parts gives

$$16\pi\, m_{ADM} = \int_{\mathbb{R}^3}\left(R(g) + \tfrac{1}{4}\|\partial g\|^2 + \|\nabla\varphi\|^2 - \tfrac{1}{2}g_{kl}\partial_i g^{jk}\partial_j g^{il}\right) dx \ .$$

Now the harmonic coordinate condition implies

$$0 = \Gamma_i \ \Leftrightarrow\ \partial_i\varphi = -g_{ij}\partial_k g^{jk} \ ,$$

and we use this to rewrite and estimate the final term

$$g_{kl}\partial_i g^{jk}\partial_j g^{il} = (g_{kl}-\delta_{kl})(\partial_i g^{jk}\partial_j g^{il} - \partial_i g^{ik}\partial_j g^{jl}) + $$
$$+ \partial_i(g^{jk}\partial_j g^{ik} - g^{ik}\partial_j g^{jk}) + g_{kl}\partial_i g^{ik}\partial_j g^{jl}$$
$$\le \tfrac{1}{5}\|\partial g\|^2 + \|\nabla\varphi\|^2 + \partial_i(g^{jk}\partial_j g^{ik} - g^{ik}\partial_j g^{jk})$$

Since g_{ij} is AF , the boundary term vanishes, and we obtain the estimate

$$16\pi\, m_{ADM} \ge \int_{\mathbb{R}^3}\left(R(g) + \tfrac{1}{8}\|\partial g\|^2 + \tfrac{1}{2}\|\nabla\varphi\|^2\right) dx\,, \tag{9}$$

so $m_{ADM} \ge 0$ and $m_{ADM} = 0$ only if g is flat. The equality case $m_{ADM} = 0$ may be considered another uniqueness theorem, and shows that the ADM mass measures the "nontriviality" of the geometry of an AF manifold with non–negative scalar curvature.

An interesting generalisation of the positive mass theorem was conjectured by R.Penrose [20] . Suppose $(\mathcal{M},g)$ satisfies the conditions of the positive mass theorem and suppose $\Sigma \longrightarrow \mathcal{M}$ is a stable, embedded, minimal 2–sphere. If $\mathcal{M}$ is a totally geodesic hypersurface (i.e. $K_{ij} \equiv 0$) of some spacetime, then Σ is the intersection of $\mathcal{M}$ and horizon of the spacetime .

Conjecture [20]:

$$4\pi(m_{ADM})^2 \ge \text{area}(\Sigma) \tag{10}$$

with equality iff $(\mathcal{M},g)$ *is the standard Schwarzschild spatial hypersurface.*

Despite the two different proofs of the positive mass theorem, there has been little progress on this conjecture.

It seems clear that we are entitled to call m_{ADM} the total mass of $(\mathcal{M},g)$. However, the linearity of the mass function in Newtonian gravity enables us to assign a mass to the separate components of a Newtonian system, by defining

$$m_N(\Omega) = \int_\Omega \mu(x)\, dx ,$$

where $\mu : \mathbb{R}^3 \to \mathbb{R}^+ \cup \{0\}$ is the mass density, and $\Omega \subset \mathbb{R}^3$. This definition generalises naturally to general relativity: if $(\mathcal{M},g)$ is an AF spacelike hypersurface with timelike unit normal vector n , then $\mu = T(n,n)$ is the local stress–energy density, and for $\Omega \subset \mathcal{M}$ we set

$$m_N(\Omega) = \int_\Omega \mu(x)\, dv_g .$$

Unfortunately this function vanishes identically for subsets of a vacuum $(T_{\alpha\beta} = 0)$ spacetime. Since physically reasonable, nontrivial, vacuum spacetimes are known to exist and these necessarily have positive ADM mass, and since we should expect a suitable *quasilocal mass* function $m(\Omega)$ to be non- zero for nonflat Ω , we must reject m_N as a candidate for a quasilocal mass. (This may have been expected, since the Einstein equations are non–linear, but the definition of m_N is based on linearity properties of Newtonian gravity).

I would now like to outline a recently proposed [21] definition of quasi- local mass, which is physically natural and poses interesting questions for geometry and physics. As motivation, recall that in Newtonian gravity, the gravitational potential u of the mass density μ satisfies

$$\begin{cases} \Delta u = 4\pi\mu & \text{in } \mathbb{R}^3 \\ u(x) \to 0 & \text{as } |x| \to \infty , \end{cases}$$

and thus we can write the total mass as

$$m = \int_{\mathbb{R}^3} \mu(x)\, dx = \frac{1}{4\pi} \int_{S(\infty)} \nabla u \cdot dS \quad ,$$

in clear analogy to the definition of the ADM mass. Similarly, we have the following formula for the Newtonian quasi–local mass

$$m_N(\Omega) = \inf \left\{ \frac{1}{4\pi} \int_{S(\infty)} \nabla u \cdot dS \text{ , where } u \in C^\infty(\mathbb{R}^3) \text{ satisfies } \Delta u = \mu \text{ in } \Omega \text{ ,} \right.$$
$$\left. \Delta u \geq 0 \text{ in } \mathbb{R}^3,\ u(x) \to 0 \text{ as } |x| \to \infty \right\} \quad ,$$

which suggests the following. Let PM be the class of Riemannian 3–manifolds,

$$\mathrm{PM} = \left\{ (\mathcal{M},g) \mid g \textit{ is a complete } \mathrm{AF} \textit{ metric of non-negative, integrable scalar curvature, } R(g) \geq 0, \textit{ and such that there are no stable minimal } 2\textit{-spheres } (\text{"horizons"}) \textit{ in } (\mathcal{M},g) \right\} .$$

Using $\Omega \subset \mathcal{M}$ to indicate Ω is an isometric subset of $(\mathcal{M},g)$, we define

$$m_B(\Omega) = \inf \left\{ m_{ADM}(\tilde{\mathcal{M}}) \mid \Omega \subset \tilde{\mathcal{M}} \in \mathrm{PM} \right\} , \tag{11}$$

for any Ω such that $\Omega \subset \mathcal{M} \in \mathrm{PM}$ and $\partial\Omega$ is connected. From the positive mass theorem we see immediately that $m_B(\Omega) \geq 0$ and $m_B(\Omega) = 0$ if $\Omega \subset \mathbb{R}^3$. Further, we have the monotonicity condition

$$m_B(\Omega_1) \leq m_B(\Omega_2) \quad \text{if } \Omega_1 \subset \Omega_2 \subset \mathcal{M} \in \mathrm{PM} \quad ,$$

which may be the best analogue we can hope for of the additivity property of the Newtonian mass. It is an interesting problem for physics to give an interpretation of this lack of additivity.

The no–horizon condition is needed to exclude a general construction of extension manifolds $(\mathcal{M}_k, g_k)$ of Ω, with $m_{ADM}(\mathcal{M}_k) \to 0$. These extensions are produced by conformally adding scalar curvature so that a horizon forms and then shrinks off, taking m_{ADM} to zero. This can easily be seen in spherically symmetric metrics.

It is an open problem to show that $m_B(\cdot)$ is a nontrivial function, i.e. that there are sets Ω with $m_B(\Omega) > 0$. However we note that any other mass function $m_C(\cdot)$ which satisfies the physically desirable condition

$$m_C \leq m_{ADM}(\mathcal{M}) \quad , \text{ for any PM extension } \Omega \subset \mathcal{M} ,$$

must satisfy also $m_C(\Omega) \leq m_B(\Omega)$, so the definition (11) is, in this sense, the best possible.

If we restrict our attention to spherically symmetric metrics, then it is easy to compute m_B. In geodesic coordinates about the origin, the spherically symmetric metric can be written

$$ds^2 = d\rho^2 + A^2(\rho)(d\theta^2 + \sin^2\theta d\varphi^2) \quad .$$

Defining the Schwarzschild mass function $m(\rho)$ by

$$m(\rho) = \frac{1}{2} A \, (1 - A'^2) \quad ,$$

the scalar curvature is determined by $m'(\rho)$;

$$m'(\rho) = \frac{1}{4} R(\rho) \, A^2 \, A' \quad . \tag{12}$$

The no–horizon condition is then $A'(\rho) > 0$, and

$$m(\rho) \longrightarrow m_{ADM} \text{ as } \rho \longrightarrow \infty ,$$

so we see that $m(\rho_0)$ is the spherical quasi–local mass of the domain $\Omega_{\rho_0} = \{\rho < \rho_0\}$. Observing that m_B is realised here by the Schwarzschild metric with mass $m(\rho_0)$, we are lead to

Conjecture: *The infimum* $m_B(\Omega)$ *is realised by a metric satisfying the static vacuum equations* (5) *outside* Ω *, and which is* $C^{0,1}$ *across* $\partial\Omega$ *, perhaps with non–negative distributional scalar curvature along* $\partial\Omega$ *.*

This static–metric conjecture can be (heuristically) justified by a pure physics argument: the extremal metric should be vacuum, since the any stress- energy can be removed by conformal change, which decreases the ADM mass. Any dynamical freedom (gravitational waves) should also increase the total mass, so the extremal metric should have a timelike isometry. Since we are assuming the extrinsic curvature K_{ij} vanishes on Ω , the extremal metric should also have $K_{ij} \equiv 0$, which implies static.

The Einstein Yang–Mills equations

Birkhoff's theorem and its generalisation [22] show there is no dynamical freedom for the spherically symmetric Einstein equations with either vacuum or Maxwell matter terms. This is not true for other matter models – Christodoulou [23] has extensively analysed the hyperbolic partial differential equations that result if the stress–energy tensor arises from pressure–free perfect fluid ("dust"), or from a massless scalar field. Although these two systems have different behaviour (for example, dust solutions can develop naked singularities, whereas it seems that the massless scalar field solutions do not), one property they have in common is the absence of non–trivial regular time–independent solutions.

In [24] we analysed the spherically symmetric Einstein Yang–Mills equations with gauge group $\mathcal{G} = SU(2)$ and showed numerically that the equations have solutions which are static, asymptotically flat, everywhere non–singular and topologically $\mathbb{R}^3 \times \mathbb{R}$. This is in contrast to the above–mentioned matter models, and is interesting for many other reasons.

Although the Yang–Mills fields are of fundamental importance in particle physics, it is known that they are "dispersive" – for finite–energy solution of the YM equations in $\mathbb{R}^{3,1}$, the energy in any bounded region in $\mathbb{R}^3$ must decay to zero [25]. The proof of this is based on the conformal invariance of the 3+1 dimensional YM equations, in particular, the invariance under the conformal dilations with generating vector $X = r\partial_r + t\partial_t$. We note that this dilational invariance fails for the EYM equations, since the mass scales as a length. The usual method for circumventing this non-existence result in $\mathbb{R}^{3,1}$ is to introduce a Higgs field – a scalar field coupled to the YM connection – and then to impose topologically non–trivial boundary conditions. Although this method is natural and attractive mathematically, it has the physical disadvantage that, unlike the YM field, the Higgs field has not been observed experimentally.

The prototypical solution of the YM–Higgs equations is the Bogomolny–Prasad–Sommerfield (BPS) monopole [26]. With generators τ_i, $i = 1,2,3$ of the gauge Lie algebra SU(2), connection $A(x)$ and Higgs field $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ valued in the adjoint representation of SU(2), the BPS monopole is given in the "canonical" gauge [27] by

$$A(x) = \left[\frac{1}{\sinh(r)} - \frac{1}{r}\right] \epsilon_{ijk}\, \tau_i \hat{x}_j\, dx_k$$

$$\Phi(x) = \left[\frac{1}{r} - \frac{1}{\tanh(r)}\right] \hat{x}_i \tau_i ,$$

where $\hat{x}_i = x_i/r$, $r = |x|$, $x \in \mathbb{R}^3$. To see that this connection is asymptotically U(1)–valued, we make the gauge change $A \rightarrow \tilde{A} = \; = u^{-1}du + u^{-1}Au$, where $u = S_3(\varphi)S_2(\theta)$, $S_i(\lambda) = \exp(\lambda\tau_i)$ and θ, φ are polar coordinates. In this "abelian" gauge [27] ,

$$\tilde{A} = \frac{r}{\sinh(r)}\left[\tau_2 d\theta - \tau_1 \sin\theta d\varphi\right] + \cos\theta\, \tau_3\, d\varphi ,$$

showing that the connection is asymptotic to the charge 1 Dirac U(1) magnetic monopole, with the non–triviality of the U(1) subbundle reflected in the fact that the gauge transformation is defined only on $\left[S^2 - \{\varphi{=}\pi,\ 0 \leq \theta \leq \pi\}\right] \times \mathbb{R}^+$. We note that the charge k Dirac monopole, $A_k = k\cos\theta\, \tau_3\, d\varphi$, serves as a source for the Reissner–Nordström metric with charge $e = k$.

We say that a connection is spherically symmetric if SO(3) (or SU(2)) acts on the principal bundle, leaving the connection invariant, and such that the projected orbits of the action on the base manifold are generically S^2 . It can then be shown [28] that in the abelian gauge, the general spherically symmetric SU(2) connection on $\mathbb{R}^{3,1}$ is:

$$A = a\tau_3 dt + b\tau_3 dr + (c\tau_1 + d\tau_2)d\theta + (c\tau_2 - d\tau_1)\sin\theta\, d\varphi + \cos\theta\, d\varphi$$

where the symmetry group is also SU(2) and a, b, c, d depend on (r,t) . We can use the remaining U(1) gauge freedom to impose the radial gauge condition $b(r,t) = 0$, and we will assume that $a(r,t) = 0$ ('t Hooft–Polyakov ansatz) .

The general static spherically symmetric, globally nonsingular metric can be written (with $m = m(r)$, $\delta = \delta(r)$)

$$ds^2 = -\left[1 - \frac{2m}{r}\right]e^{-2\delta}dt^2 + \left[1 - \frac{2m}{r}\right]^{-1}dr^2 + r^2 d\Omega^2 , \qquad (13)$$

and the static, spherically symmetric EYM equations reduce to the system

$$\text{(i)} \qquad \left[e^{-\delta}\left[1-\frac{2m}{r}\right]w'\right]' + e^{-\delta}\left[1-w^2\right]w/r = 0$$

$$\text{(ii)} \qquad \delta' = -2\,\frac{w'^2}{r} \tag{14}$$

$$\text{(iii)} \qquad m' = \left[1-\frac{2m}{r}\right]w'^2 + \frac{(1-w^2)^2}{2r^2},$$

where the YM connection is given by setting $d = w$, $c = 0$. The equations (14 (ii,iii)) show that the metric functions δ, m are determined by the YM function w, which satisfies the YM equation (14 (i)). In fact, if we use (14 (ii,iii)) to define (m,δ) from w, the YM equation can be obtained as the Euler–Lagrange equation of the (nonlocal) total mass functional

$$M(w) = \int_0^\infty e^{-\delta}\left[w'^2 + \frac{(1-w^2)^2}{2r^2}\right]dr .$$

The mass equation (iii) is just (12) with stress–energy density

$$T_{00} = \frac{1}{2}\|B\|^2 = \left[1-\frac{2m}{r}\right]\left[\frac{w'}{r}\right]^2 + \frac{1}{2}\left[\frac{1-w^2}{r^2}\right]^2 .$$

The metric is asymptotically flat if we impose the boundary condition

$$M(w) = \lim_{r\to\infty} m(r) < \infty , \tag{15}$$

and the solution is regular across the origin if we require that the stress–energy is bounded,

$$\lim_{r\to 0} T_{00}(r) < \infty .$$

This leads to the boundary conditions

$$m(0) = 0$$
$$w(r) = 1 + O(r^2) \quad \text{as } r \to 0 , \qquad (16)$$

which we used as the basis of a "shooting method" search for numerical solutions [24] . (We note that the boundary condition (16) ensures that the connection can be gauge transformed into a smooth connection on a bundle extending smoothly across the origin). Numerically, a discrete family of solutions satisfying the asymptotic boundary condition (15) were found, indexed by the number of zeroes of $w(r)$. In Figures 1,2 we show the YM function $w(r)$ and the density T_{00} for the 4–zero solution. The density is concentrated in the region $r < 1$ and decays rapidly (but polynomially). In the farfield region $r > 1000$, the solution approximates the Schwarzschild solution since $w \simeq 1$, and the mass is $M_4 \simeq 0.999236$. (I do not understand why $M_k \sim 1$) . In the near field region, $1 < r < 1000$, the solution $w \sim 0$ indicates the curvature is approximately that of a charge 1 Dirac monopole, and the metric is approximately Reissner–Nordström. We quantify this by introducing the magnetic charge function $g^2(r)$ by

$$e^{-\delta(r)}\left(1 - \frac{2m(r)}{r}\right) = 1 - \frac{2M}{r} + \frac{g^2}{r^2} \quad ,$$

and Figure 3 shows that, as expected, the RN charge g^2 agrees with the Dirac magnetic charge 1 in the near field region. Since the total mass $M \simeq 1$, the metric in the near field region in fact approximates the extremal RN metric, (2) with $m = e\ (= g)$.

The high–density core and asymptotic flatness suggest that these solutions are classical particles, but unfortunately this is probably not a valid interpretation, since we expect the time–dependent solutions to be unstable. This can be seen by considering the solutions as resulting from a balance between the attractive gravitational force and the repulsive YM force, typified by collapse to a black hole on one hand, and dispersion of YM radiation on the other. This unstable gravitational collapse/YM dispersion balance has been observed numerically.

There are clearly many questions about the EYM equations which need to be answered, starting with a proof of the existence of the numerical solutions found in [24] . The static spherically symmetric EYM equations without the 't Hooft–Polyakov ansatz (i.e. nonvanishing electric field) are also very interesting, as we may hope that the electric field helps to stabilise the solution. The numerical evidence for the existence of asymptotically flat solutions here is inconclusive, but does show that the equations with electric field behave quite differently.

The observation that the solutions approximate the extremal RN metric in the near–field region raises the intriguing possibility of the existence of superposition solutions of the static EYM , approximating the Majumdar- Papapetrou metrics [22,29]

$$ds^2 = -U^{-2}\, dt^2 + U^2 |dx|^2$$

where $|dx|^2$ is the standard metric on $\mathbb{R}^3$, and $U(x)$ satisfies $\Delta U = 0$. If $p_1,\ldots,p_k \in \mathbb{R}^3$ and $r_k(x) = |x-x_k|$, then the MP metric with

$$U(x) = 1 + \sum_{i=1}^{k} \frac{m_i}{r_i} \ ,$$

represents the superposition of extremal Reissner-Nordström metrics with charges $e_i = m_i$, positioned at p_i [29]. This would be analogous to the superposition of multi–monopole solutions of the Bogomolny-YMH equations [30]. Since the positions of the extremal RN sources in the MP metric are arbitrary, we may expect that the EYM particles move freely as long as their separation distances fall below the near–field radius. Beyond this radius, we should not expect static solutions, since the RN charge decays and the gravitational attractions predominates (this is assuming some mechanism can be found which provides stability) . However, unless an algebraic construction can be found, analogous to the constructions for multimonopole BYMH solutions, it seems unlikely that these solutions could be found by other that numerical computation. This applies in particular to the time–dependent evolution, which seems completely intractable to current analytic techniques.

References、

[1] A. Einstein, Ann. Phys. **49** , 769–822 (1916)

[2] R. Schoen and S.–T.Yau, Comm.Math.Phys. **65**, 45–76 (1979); **79**, 231–260 (1981)

[3] E. Witten, Comm. Math. Phys. **80**, 381–402 (1981)

[4] K. Schwarzschild, Sitz. Preuss. Akad. Wiss. 189, 424 (1916)

[5] S. Hawking and D.Ellis, *Large Scale Structure of Spacetime*, Cambridge UP, 1973

[6] H. Reissner, Ann. Physik **50**, 106 (1916)

G. Nordström, Proc. Kon. Ned. Akad. Wet. **20**, 1238 (1918)

[7] G. Birkhoff, *Relativity and Modern Physics*, Harvard UP (1923)

[8] W. Israel, Comm. Math. Phys. **8**, 245–260 (1968)

D. Robinson, Gen. Rel. Grav. **4**, 53–78 (1973)

G. Bunting and A.K.M. Masood–ul–Alam, Gen. Rel. Grav. **19**, 147–154 (1987)

[9] R. Bach and H. Weyl, Math. Zeit. **13** ,134–145 (1922)

[10] A.K.M. Massood–ul–Alam, Class. Quant. Grav. **5**, 409–422 (1988)

[11] B. Carter, in *Black Holes: 1972 Les Houches*, editors B.,C. De Witt

D. Robinson, Phys. Rev. Lett. **34**, 905–906 (1975)

G. Bunting, PhD. Thesis, Univ. New England, 1983

P. Mazur, J. Phys. A **15** , 3173–3175 (1982)

[12] S.–T. Yau, in Ann. Math. Studies 102, Princeton 1981

[13] J. Eschenburg, J. Diff. Geom. **27** , 477–492 (1988)

[14] G. Galloway, J. Diff. Geom. **29** , 373–388 (1989) and in *Mathematical Relativity*, CMA Proceedings Vol. 19, 1988

R. Newman, to appear in J. Diff. Geom.

[15] J. Cheeger and D. Gromoll, J. Diff. Geom. **6** , 141–151 (1971)

[16] J.Beem, P.Ehrlich, S.Markvorsen and G.Galloway, J.Diff.Geom. **22**, 29–42 (1985)

[17] R. Arnowitt. S. Deser and C. Misner, Phys. Rev. **122** , 997–1006 (1961)

[18] R.Bartnik, Comm.Math.Phys. **94**, 155–174(1984), Acta Math. **161**; 145–181(1988)

[19] R. Bartnik, Comm. Pure Appl. Math. **39** , 661–693 (1986)

[20] R. Penrose, Ann. N.Y. Acad. Sc. **224** , 125–134 (1973)

[21] R. Bartnik, Phys. Rev. Lett. **62** , 2346–2348 (1989)

[22] D. Kramer, H. Stephani, M. MacCallum and E. Herlt, *Exact solutions of Einstein's field equations* , Cambridge 1980 .

[23] D. Christodoulou, Comm. Math. Phys. **93** 171–195 (1984); **105** 337–361; **106**, 587–621; **109** 591–611; **109** 613–647 (1987)

[24] R. Bartnik and J. MacKinnon, Phys. Rev. Lett. **61** 615–624 (1988)

[25] S. Coleman, in *New Phenomenon in Subnuclear Physics,* ed. A. Zichichi, Plenum 1975

S. Deser, Phys. Lett. **64 B**, 463–465 (1976)

R. Weder, Comm. Math. Phys. **57**, 161–164 (1977)

[26] A. Jaffe and C. Taubes, *Vortices and Monopoles* , Birkhäuser Boston 1980.

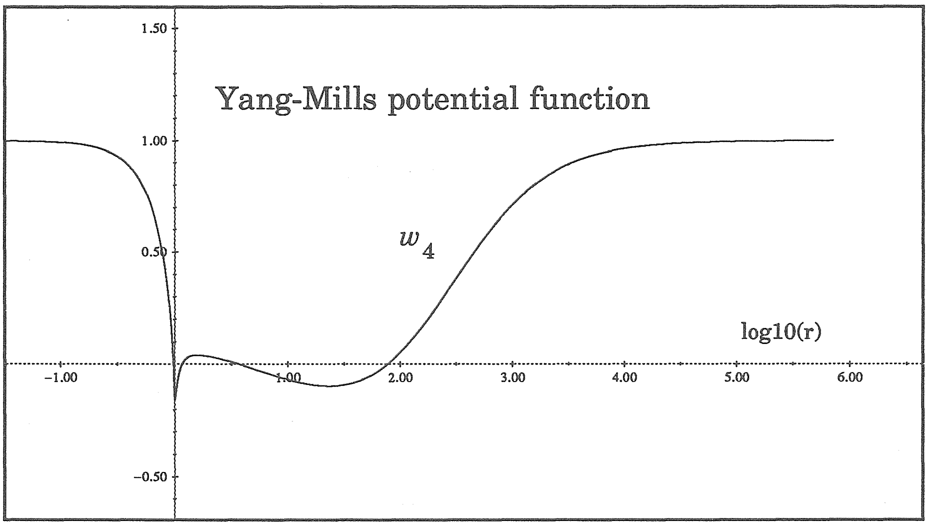
[27] N. Manton, Ann. Physics **132** 108–120 (1981)

[28] E. Witten, Phys. Rev. Lett. **38** 121–124 (1977)

J. Harnad, S. Shnider and L. Vinet, J. Math. Phys. **21** 2719–2724 (1980)

[29] J. Hartle and S. Hawking, Comm. Math. Phys. **26**, 87-101 (1972)

[30] M. Atiyah and N. Hitchin, *The Geometry and Dynamics of Magnetic Monopoles* , Princeton, 1988.



Yang-Mills potential function
1.50
1.00
0.50
-0.50
w_4
log10(r)
-1.00
1.00
2.00
3.00
4.00
5.00
6.00

Figure 1.

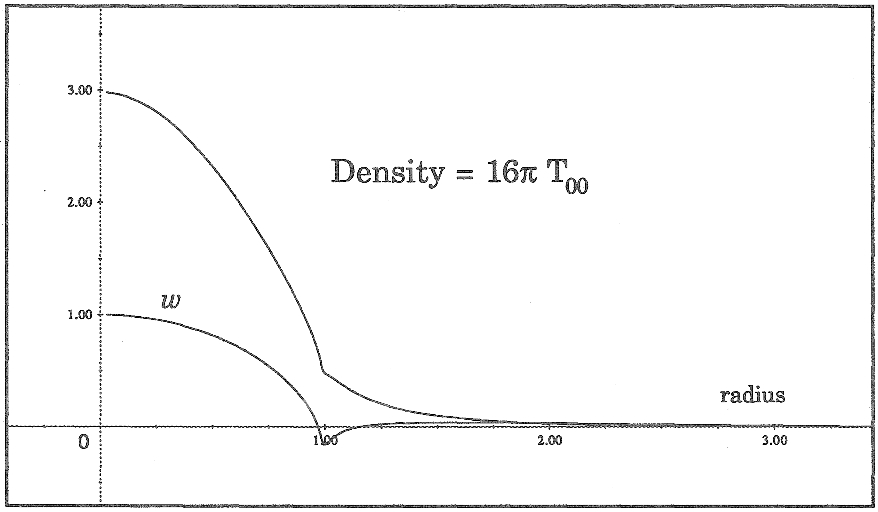


Figure 2.

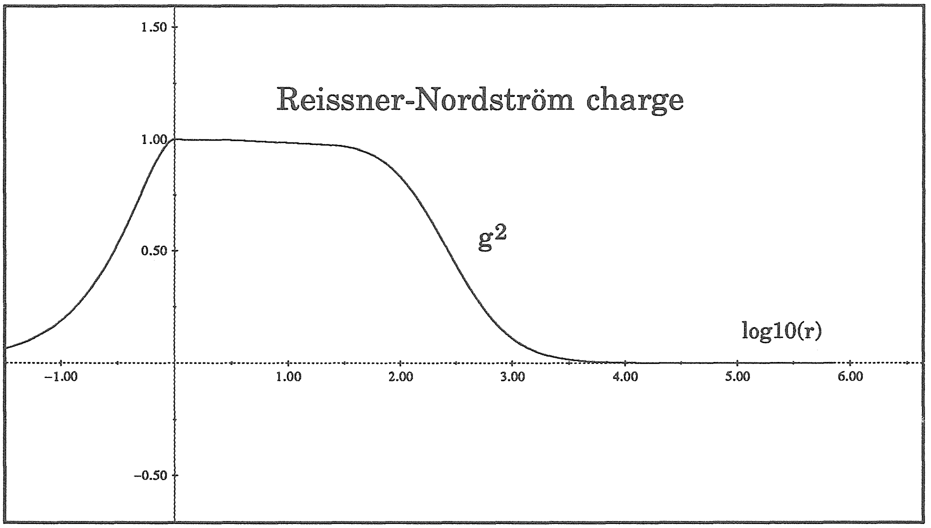

1.50
1.00
0.50
-0.50
-1.00
1.00
2.00
3.00
4.00
5.00
6.00
Reissner-Nordström charge
g^2
log10(r)

Figure 3.

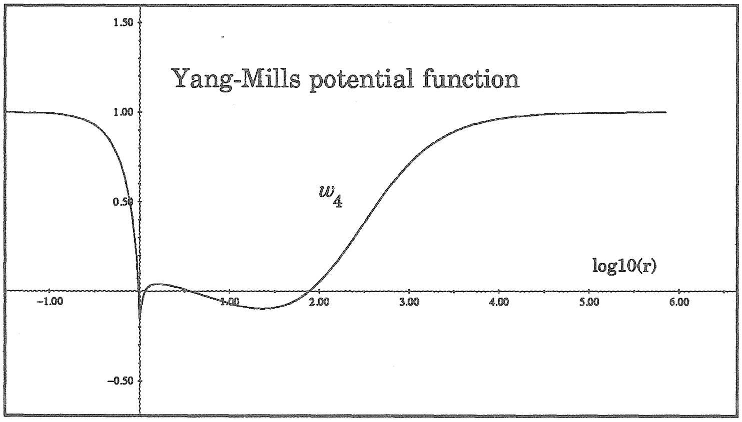

Yang-Mills potential function
w_4
log10(r)
1.50
1.00
0.50
-0.50
-1.00
1.00
2.00
3.00
4.00
5.00
6.00

Figure 1.

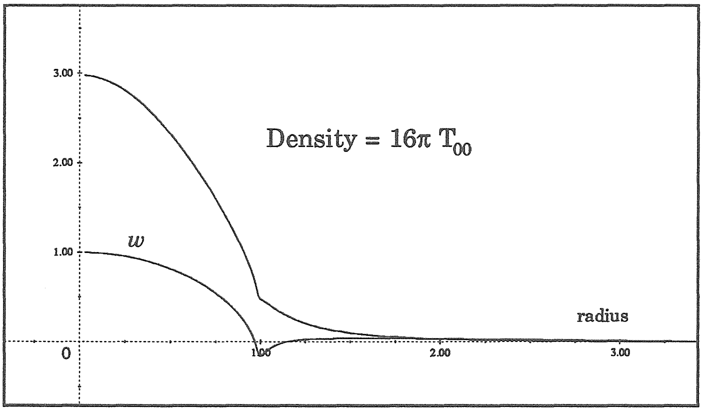


Figure 2.

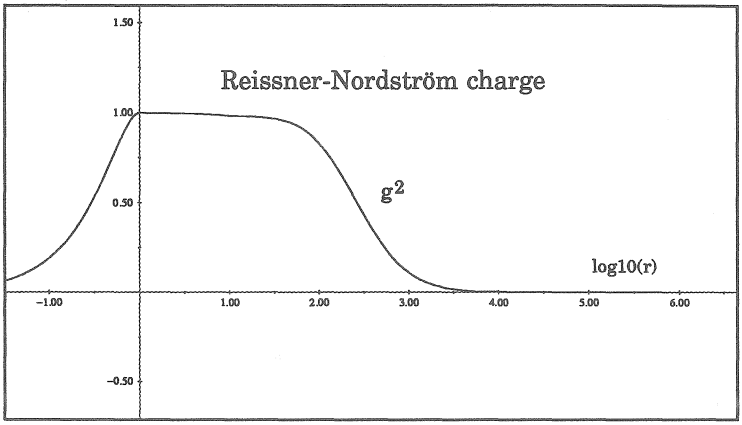

Reissner-Nordström charge
g2
log10(r)
1.50
1.00
0.50
-0.50
-1.00
1.00
2.00
3.00
4.00
5.00
6.00

Figure 3.